A&A manuscript no.
(will be inserted by hand later)

Your thesaurus codes are:
02(12.07.1; 03.13.1; 03.13.4)

ASTRONOMY AND ASTROPHYSICS
June 8, 1999

A fast direct method of mass reconstruction for gravitational lenses

M. Lombardi[1,2] **and G. Bertin**[2]

[1] European Southern Observatory, Karl-Schwarzschild Straße 2, D 85748 Garching bei München, Germany
[2] Scuola Normale Superiore, Piazza dei Cavalieri 7, I 56126 Pisa, Italy

Received 25 March 1999; accepted 17 May 1999

Abstract. Statistical analyses of observed galaxy distortions are often used to reconstruct the mass distribution of an intervening cluster responsible for gravitational lensing. In current projects, distortions of thousands of source galaxies have to be handled efficiently; much larger data bases and more massive investigations are envisaged for new major observational initiatives. In this article we present an efficient mass reconstruction procedure, a direct method that solves a variational principle noted in an earlier paper, which, for rectangular fields, turns out to reduce the relevant execution time by a factor from 100 to 1000 with respect to the fastest methods currently used, so that for grid numbers $N = 400$ the required CPU time on a good workstation can be kept within the order of 1 second. The acquired speed also opens the way to some long-term projects based on simulated observations (addressing statistical or cosmological questions) that would be, at present, practically not viable for intrinsically slow reconstruction methods.

Key words: cosmology: gravitational lensing – methods: analytical – methods: numerical

1. Introduction

In the context of weak or statistical lensing, the problem of the determination of the dimensionless mass density distribution $\kappa(\boldsymbol{\theta})$ from a map of the reduced shear $g(\boldsymbol{\theta})$ has been considered in detail by various authors, using either simulations or analytical calculations (e.g., Bartelmann 1995, Schneider 1995, Seitz & Schneider 1996, Squires & Kaiser 1996, Lombardi & Bertin 1998a, 1998b).

The mass inversion is usually performed starting from the vector field $\tilde{\boldsymbol{u}}(\boldsymbol{\theta})$ defined in terms of the *measured* reduced shear $g(\boldsymbol{\theta})$ (Kaiser 1995). In the ideal case where the *measured* shear $g(\boldsymbol{\theta})$ is just the *true* shear $g_0(\boldsymbol{\theta})$, the vector field $\tilde{\boldsymbol{u}}_0(\boldsymbol{\theta})$ can be shown to satisfy the relation

$$\tilde{\boldsymbol{u}}_0(\boldsymbol{\theta}) = \nabla \ln\bigl[1 - \kappa_0(\boldsymbol{\theta})\bigr] = \nabla\tilde{\kappa}_0(\boldsymbol{\theta}) \,, \tag{1}$$

Send offprint requests to: M. Lombardi
Correspondence to: lombardi@sns.it

where κ_0 is the true dimensionless mass map and $\tilde{\kappa}_0(\boldsymbol{\theta}) = \ln\bigl[1 - \kappa_0(\boldsymbol{\theta})\bigr]$. However, because of statistical and measurement errors, $\tilde{\boldsymbol{u}}$ is not necessarily curl-free, and thus κ can be determined only approximately. In a separate paper (Lombardi & Bertin 1998b) we have shown that

- The statistical errors on $\kappa(\boldsymbol{\theta})$ are minimized if this function is calculated as

$$\tilde{\kappa}(\boldsymbol{\theta}) = \bar{\kappa} + \int_\Omega \boldsymbol{H}^{\mathrm{SS}}(\boldsymbol{\theta},\boldsymbol{\theta}') \cdot \tilde{\boldsymbol{u}}(\boldsymbol{\theta}')\, \mathrm{d}^2\theta' \,, \tag{2}$$

  where $\bar{\kappa}$ is a constant (introduced to take into account the *sheet invariance*), $\boldsymbol{H}^{\mathrm{SS}}(\boldsymbol{\theta},\boldsymbol{\theta}')$ is the noise-filtering kernel (Seitz & Schneider 1996), and Ω is the field of observation.
- The same mass map can be obtained by solving the equations

$$\nabla^2\tilde{\kappa} = \nabla\cdot\tilde{\boldsymbol{u}} \,, \tag{3}$$

$$\nabla\tilde{\kappa}\cdot\boldsymbol{n} = \tilde{\boldsymbol{u}}\cdot\boldsymbol{n} \qquad \text{on } \partial\Omega \,, \tag{4}$$

  where $\boldsymbol{n}$ is the unit vector perpendicular to the boundary $\partial\Omega$ of the field of observation Ω. Hence, the kernel $\boldsymbol{H}^{\mathrm{SS}}(\boldsymbol{\theta},\boldsymbol{\theta}')$ can be identified as the Green function of this Neumann boundary problem.
- Equations (3) and (4) are precisely the Euler equations associated with the functional

$$S = \frac{1}{2}\int_\Omega \bigl\|\nabla\tilde{\kappa}(\boldsymbol{\theta}) - \tilde{\boldsymbol{u}}(\boldsymbol{\theta})\bigr\|^2 \mathrm{d}^2\theta \,. \tag{5}$$

  In other words, the functional S is minimized when $\tilde{\kappa}(\boldsymbol{\theta})$ is calculated using Eq. (2) or, equivalently, by solving Eqs. (3) and (4).

To these three formulations of the mass inversion problem there correspond three practical methods to calculate $\kappa(\boldsymbol{\theta})$ from a given set of data.

- The first method considered is based on a direct calculation of the kernel $\boldsymbol{H}^{\mathrm{SS}}$ (Seitz & Schneider 1996). Once this kernel has been calculated for a given field Ω, the mass inversion is straightforward (note that the kernel $\boldsymbol{H}^{\mathrm{SS}}$ depends on the field of observation). A problem with this method is that a calculation of

$\boldsymbol{H}^{\rm SS}$ is expensive in terms of memory requirements and computation time. In fact, in order to compute κ on a square grid of $N \times N$ points, $\boldsymbol{H}^{\rm SS}$ must be calculated on a multidimensional grid of $N \times N \times N \times N$ points. Moreover, $2N^4$ multiplications are needed to evaluate Eq. (2), and thus the method is of order $\mathcal{O}\big(N^4\big)$. Because of the large memory needed to allocate $\boldsymbol{H}^{\rm SS}$, calculations can be performed only with a limited value of N (typically $N \sim 50$).

– The introduction of a method that directly solves the Neumann problem allows one to go beyond many of the limitations of the $\boldsymbol{H}^{\rm SS}$ method. Equations (3) and (4) can be solved using an *over-relaxation method* (Seitz & Schneider 1998). In this case $\tilde\kappa(\boldsymbol{\theta})$ is calculated directly, and thus we need to allocate only $N \times N$ real numbers. Moreover, the method can be applied without difficulties to "strange" geometries Ω (while the previous method is straightforward only when applied to rectangular or circular fields). The over-relaxation method is quicker than the kernel method, being approximately of order $\mathcal{O}\big(N^3\big)$.

– A *direct* method to minimize the functional (5) will be presented in this paper. As we will see, this method has several advantages and turns out to be very efficient from a computational point of view, being of order $\mathcal{O}\big(N^2 \log N\big)$. Moreover, it is extremely easy to implement (two implementations for rectangular fields Ω written in `C` and in `IDL` are freely available on request).

We should stress that, as proved in an earlier paper (Lombardi & Bertin 1998b), the three formulations are mathematically equivalent. Thus it would not be surprising to find that proper numerical implementations perform, for large values of the grid number N, in a similar manner as far as accuracy and reliability are concerned. In practice, for finite values of N, the third method turns out to be characterized by small errors, often smaller than those associated with the other two procedures.

2. A direct method to solve the variational principle

Direct methods in variational problems are well-known especially in applied mathematics (see, e.g., Gelfand & Fomin 1963). Suppose that one can find a *complete* set of functions $\{f_\alpha\}$ on the domain Ω (the full definition of "complete" will be given below), so that any function on Ω can be represented as a linear combination of the form

$$\tilde\kappa(\boldsymbol{\theta}) = \sum_{\alpha=1}^{\infty} c_\alpha f_\alpha(\boldsymbol{\theta}) \; . \tag{6}$$

More precisely, we assume that for any function $\tilde\kappa(\boldsymbol{\theta})$, there is a choice for the coefficients $\{c_\alpha\}$ such that

$$\int_\Omega \left\| \nabla\tilde\kappa(\boldsymbol{\theta}) - \sum_{\alpha=1}^{\infty} c_\alpha \nabla f_\alpha(\boldsymbol{\theta}) \right\|^2 {\rm d}^2\theta = 0 \; . \tag{7}$$

Let us now introduce a sequence of trial mass maps

$$\tilde\kappa^{[n]}(\boldsymbol{\theta}) = \sum_{\alpha=1}^{n} c_\alpha^{[n]} f_\alpha(\boldsymbol{\theta}) \; . \tag{8}$$

We further require that the function $\tilde\kappa^{[n]}$ minimizes the functional S: in other words, $c_1^{[n]}, c_2^{[n]}, \ldots, c_n^{[n]}$ are chosen so that the functional S has minimum value. This obviously happens when

$$\frac{\partial S}{\partial c_\alpha^{[n]}} = 0 \qquad \text{for } \alpha = 1, 2, \ldots, n \; . \tag{9}$$

Solving this set of n equations, we obtain the n coefficients $c_\alpha^{[n]}$, and thus the function $\tilde\kappa^{[n]}$. By repeating this operation for a sequence of values of n, we find a sequence of functions $\tilde\kappa^{[n]}$. These functions, under suitable assumptions (verified in our problem), have the following properties (see Gelfand & Fomin 1963 for a detailed discussion): (i) Let us call $S^{[n]}$ the value of S when $\tilde\kappa$ is replaced by the function $\tilde\kappa^{[n]}$. Then, obviously, the sequence $S^{[n]}$ is not increasing. (ii) If the set $\{f_\alpha\}$ is complete, then the functions $\tilde\kappa^{[n]}$ converge to the solution $\tilde\kappa$ of the problem. This method thus provides a way to obtain the function $\tilde\kappa(\boldsymbol{\theta})$ with desired accuracy.

The method described here can be easily applied to our problem. In fact, by expanding $\tilde\kappa^{[n]}(\boldsymbol{\theta})$ as in Eq. (8), we find

$$\frac{\partial S}{\partial c_\alpha^{[n]}} = \int_\Omega \nabla f_\alpha(\boldsymbol{\theta}) \cdot \left[\sum_{\beta=1}^{n} c_\beta^{[n]} \nabla f_\beta(\boldsymbol{\theta}) - \tilde{\boldsymbol{u}}(\boldsymbol{\theta}) \right] {\rm d}^2\theta = 0 \; . \tag{10}$$

The previous equation, for $\alpha = 1, 2, \ldots, n$, represents a linear system of n equations for the n variables $\big\{c_\alpha^{[n]}\big\}$. Its solution is thus the set of coefficients to be used in Eq. (8).

However, we note that care must be taken in the choice of a *complete* set of functions. Let us define, for the purpose, the product $\langle \boldsymbol{v}, \boldsymbol{w} \rangle$ between two generic vector fields $\boldsymbol{v}(\boldsymbol{\theta})$ and $\boldsymbol{w}(\boldsymbol{\theta})$ as

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \int_\Omega \boldsymbol{v}(\boldsymbol{\theta}) \cdot \boldsymbol{w}(\boldsymbol{\theta}) {\rm d}^2\theta \; . \tag{11}$$

As our problem involves $\nabla\tilde\kappa$, the completeness has to be referred to the set of the gradients. In other words, the set $\big\{f_\alpha\big\}$ is complete if

$$\langle \nabla f_\alpha, \nabla\tilde\kappa \rangle = 0 \tag{12}$$

for every α implies $\tilde\kappa(\boldsymbol{\theta}) =$ constant. It is easy to show that this condition is equivalent to Eq. (7).

The direct method can be further simplified if a set of functions $\{f_\alpha\}$ can be taken to satisfy a suitable orthonormality condition, so that the gradients of the functions $\{f_\alpha\}$ satisfy

$$\langle \nabla f_\alpha, \nabla f_\beta \rangle = \delta_{\alpha\beta} \; , \tag{13}$$

where $\delta_{\alpha\beta} = 1$ for $\alpha = \beta$ and 0 otherwise. Then Eq. (10) can be rewritten simply as

$$c_\alpha^{[n]} = \langle \nabla f_\alpha, \tilde{\boldsymbol{u}} \rangle \ . \tag{14}$$

Thus, with the use of an orthonormal set of functions we have secured two important advantages: (i) The linear system (10) has been diagonalized, so that its solution is trivial. (ii) The coefficients $c_\alpha^{[n]}$ no longer depend on n: that is, the coefficients of the *exact* solution are given by $c_\alpha = \langle \nabla f_\alpha, \tilde{\boldsymbol{u}} \rangle$.

Because of these advantages, whenever possible an orthonormal set of functions should be used. We note, however, that the orthonormality condition (13) depends on the field of observation Ω. Even if the existence of an orthonormal set of functions is always guaranteed by the spectral theory for the Laplace operator (see Brezis 1987), for "strange" geometries, it may be *non trivial* to find a complete orthonormal set of functions. In such cases, we need to solve the linear system (10).

The direct method described above has several advantages with respect to the "kernel" method and to the over-relaxation method: (i) The method is fast in the case where an orthonormal set of functions can be found. In fact, we need only to evaluate one integral for each coefficient c_α that we want to calculate. (ii) The method does not require a large amount of memory: we need to retain only the n values of the coefficients c_α. (iii) The precision of the inversion is driven in a natural way by the value of n. Typically, the larger α is, the smaller the length scale of f_α (see below). (iv) In some cases, the decomposition of the mass density $\tilde{\kappa}(\boldsymbol{\theta})$ in terms of the functions f_α can be useful.

3. Rectangular fields

When the field Ω is rectangular, an orthonormal set of functions can be written easily. Here we consider the special case when Ω is a square of length π (in some suitable units); any rectangular field can be handled in a similar manner. In the case considered, an orthonormal set of functions is given by

$$f_{\alpha\beta}(\boldsymbol{\theta}) = n_{\alpha\beta} \cos\alpha\theta_1 \cos\beta\theta_2 \ , \tag{15}$$

with $(\alpha, \beta) \in \mathbb{N}^2 \setminus (0,0)$. The normalization $n_{\alpha\beta}$ is defined as

$$n_{\alpha\beta} = \begin{cases} \dfrac{\sqrt{2}}{\pi\sqrt{\alpha^2+\beta^2}} & \text{for } \alpha = 0 \text{ or } \beta = 0 \ , \\ \dfrac{2}{\pi\sqrt{\alpha^2+\beta^2}} & \text{otherwise} \ . \end{cases} \tag{16}$$

The function f_{00} is not defined. Note that here we use two indices for the set. Cosines must be used in order to have a *complete* set (see Eqs. (7), (12), and Appendix A). Our problem is solved in terms of the coefficients $c_{\alpha\beta}$:

$$c_{\alpha\beta} = -n_{\alpha\beta} \int_\Omega \big[\alpha\tilde{u}_1(\boldsymbol{\theta}) \sin\alpha\theta_1 \cos\beta\theta_2 + \beta\tilde{u}_2(\boldsymbol{\theta}) \cos\alpha\theta_1 \sin\beta\theta_2\big] \mathrm{d}^2\theta \ , \tag{17}$$

$$\tilde{\kappa}(\boldsymbol{\theta}) = \sum_{\alpha\beta} c_{\alpha\beta} n_{\alpha\beta} \cos\alpha\theta_1 \cos\beta\theta_2 \ . \tag{18}$$

We now observe that the particular choice of the orthonormal set $\{f_{\alpha\beta}\}$ allows us to use fast Fourier transform (FFT) techniques to evaluate Eqs. (17) and (18). The use of FFT makes the direct method very efficient: in particular the method becomes of order $\mathcal{O}\big(N^2 \log N\big)$. Moreover, several optimized FFT libraries are available.

The optimal truncation for the series (18) is determined by the adopted grid numbers: for a grid of $N \times M$ points, α should run from 0 to $N-1$, and β from 0 to $M-1$ (this is standard practice for FFT libraries).

4. Performance

Our method has been implemented in `C` and in `IDL`. The `C` version uses the library `FFTW` ("Fastest Fourier Transform in the West," version 2.0.1) to perform discrete Fourier transforms (DFT). This library, written by Matteo Frigo and Steven G. Johnson, is considered the quickest DFT library publicly available. The performance of our direct method is compared with that of the over-relaxation method, also implemented in `C`. The procedure used in the tests is summarized in the following points:

1. A simple model for the dimensionless mass distribution has been chosen. Then the mass distribution $\tilde{\kappa}_0$ is calculated on a grid of $N \times N$ points.
2. The associated field $\tilde{\boldsymbol{u}}_0$ is calculated on the same grid using a 3-point Lagrangian interpolation in order to numerically evaluate the derivatives that are needed.
3. Noise is added to the vector field $\tilde{\boldsymbol{u}}_0$ using an analytical model for the noise derived earlier (Lombardi & Bertin 1998b). In practice, the various Fourier components of the noise are added using a suitable model for the power spectrum.
4. The resulting noisy $\tilde{\boldsymbol{u}}$ map is inverted using the over-relaxation method and the present direct method. The two dimensionless mass maps obtained are then compared. Moreover, the inversion times are recorded.

The results obtained in the tests are the following:

- The two mass densities obtained are consistent with each other.
- Because of the set of functions used, the errors produced by the direct method are larger on the boundary of the field. For this reason, we suggest that a one pixel strip around the field should be discarded. The area discarded is very small.
- Some tests have been performed by providing $\tilde{\boldsymbol{u}}_0$ to the inversion procedures. This allows us to compare the reconstructed mass density with the original map κ_0. From such tests we have noted that the discretization

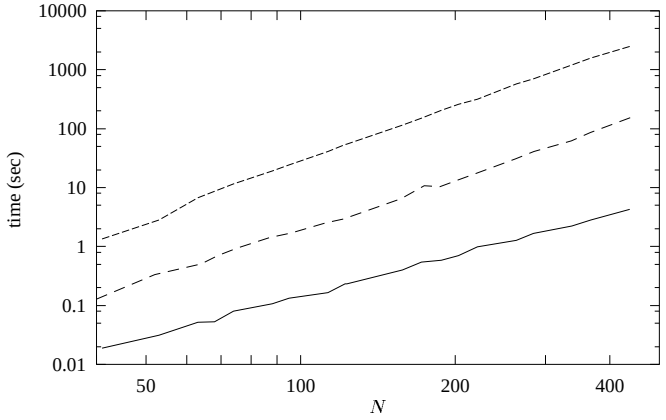

Fig. 1. Execution time per call vs. grid number N. The solid line refers to the direct method applied to "good numbers" (values $2N-1$ that can be factorized with small primes), the long-dashed refers to the direct method applied to "bad numbers" ($2N-1$ prime), and the short-dashed line to the over-relaxation method.

errors of the direct method are slightly smaller than the discretization errors of the over-relaxation method.

– The results of the two methods differ because of the *sheet invariance*: in particular, the direct method always gives a "reduced" mass map $\tilde{\kappa}$ with vanishing total mass.

Regarding the second item, we note that the error is related to the finite sampling scale of the method; the error affects only the outermost pixel because of the proper choice of the truncation (see comment at the end of Sect. 3).

The measured execution times are plotted in Fig. 1 for different values of N. These are the averaged CPU execution times for a single reconstruction on a SUN Ultra 1 workstation. From this figure it is clear that the direct method is much faster than the over-relaxation method. Here we should recall that, because of some characteristics of the `FFTW` library, the execution time of the direct method can change significantly even for neighbouring values of N. In particular, the inversion is faster when $(2N-1)$ can be factorized with small prime numbers, and is slower in other cases (see Fig. 1). For example, the execution time (on a SUN Ultra 1) changes from 2.942 to 0.232 seconds when N changes from 121 to 122. Finally, we observe that our implementation of the direct method is not optimal: in fact, with a different (non-trivial) use of FFT one might gain an additional factor of 4 on the execution time.

Besides the appealing aspects of simplicity inherent to the direct method described in this paper, we should note that gaining three orders of magnitude in CPU time will make it possible to undertake a few long-term projects of simulated observations (in particular, with the goal of a statistically sound investigation of the quality of mass reconstruction; but other objectives might be formulated, e.g. in the cosmological context) that would remain practically out of reach for other intrinsically slow reconstruction methods.

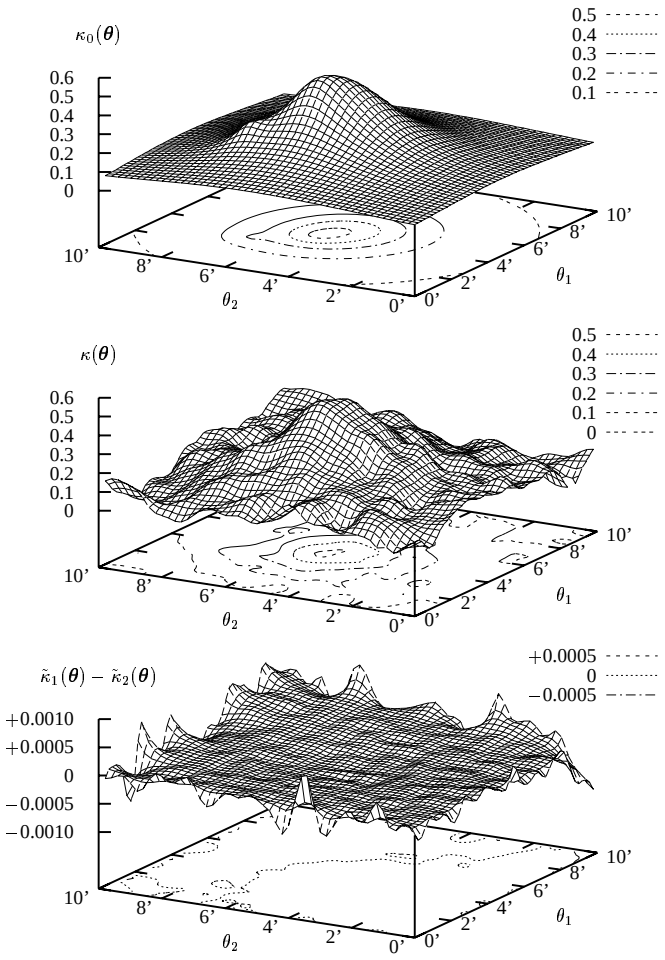

Fig. 2. A typical result of mass reconstruction; at the adopted distance for the lensing cluster, the side of the square field, 10 arcminutes, corresponds to approximately 2.88 Mpc. From top to bottom, true dimensionless mass distribution, reconstructed distribution (from the direct method), and difference between maps of the variable $\tilde{\kappa}$ derived from direct and over-relaxation methods. The very small residuals show that the two methods are practically equivalent in terms of accuracy.

5. Examples of simulated reconstructions

In addition to the reconstructions from "synthetic" data as described in the previous Section, we have performed several additional tests in order to demonstrate the reliability of our method. The tests, designed with the aim to reproduce the main features of a "real" reconstruction, have been made following a straightforward procedure (see, e.g., Lombardi & Bertin 1999 for similar simulations).

First we have generated a population of source galaxies using a pseudo-random number generator. Here a source galaxy is represented by its position and by its ellipticity

(see Eq. (11) of Lombardi & Bertin 1999, following Seitz & Schneider 1997). Positions are drawn from a homogeneous distribution (with a density of 70 galaxies per square arcmin), while ellipticities are drawn from a truncated Gaussian distribution with variance $\sigma^2 = (0.3)^2$. Sources are assumed to have all the same redshift $z_\mathrm{s} = 1.5$. Source ellipticities are then transformed into observed ellipticities. For simplicity, the observed galaxy positions are assumed to be equal to the source positions: in other words, no depletion effects are included in the simulations.

Then the calculation of the observed ellipticities has been done by referring to a cluster of galaxies placed at $z_\mathrm{d} = 0.3$ with total mass inside the 10' × 10' field $1.5 \times 10^{15}\, M_\odot$. For the purpose of introducing the lensing effects, we only need to specify the dimensionless projected mass map $\kappa(\boldsymbol{\theta})$. For simplicity, we have used a density distribution made of three symmetrical components; each component is described by the analytical model outlined by Schneider *et al.* (1992, p. 244), which, at large radii, is approximately isothermal.

By averaging the observed ellipticities, we have then obtained a map of the reduced shear $g(\boldsymbol{\theta})$ and, from that map, the vector field $\tilde{\boldsymbol{u}}(\boldsymbol{\theta})$. The mass inversion has been performed using the direct method and the over-relaxation method. The two mass distributions have been then compared.

An example of typical results is shown in Fig. 2. Here, from top to bottom, we display the original cluster mass distribution, the reconstruction obtained using the direct method, and the residuals, i.e. the difference between the reconstructed maps from the direct method and from the over-relaxation method. As the figure clearly shows, differences are mainly confined to the boundary of the field where they are found to be of the order of 0.0002, well below the statistical errors of the reconstruction. In the inner field the differences are about one order of magnitude smaller. Note also that the wavy overall appearance of the reconstructed map is normal for weak lensing reconstructions, resulting from the relatively low number of source galaxies involved (see Lombardi & Bertin 1998b for a discussion of the statistical aspects of the problem).

Acknowledgements. We thank Luigi Ambrosio and Peter Schneider for interesting discussions and suggestions. The reconstruction code uses `FFTW 2.0.1` by Matteo Frigo and Steven G. Johnson. This work has been partially supported by MURST and by ASI of Italy.

Appendix A: Completeness of $\{f_{\alpha\beta}\}$

In this appendix we will verify explicitly that the set of functions defined in Eq. (15) is complete, in the sense that Eqs. (3) and (4) can be recovered. For the purpose, we will apply the Fourier theorem (Brezis 1987). In the following, we will assume that $\tilde{\boldsymbol{u}}(\boldsymbol{\theta})$ is a *smooth* vector field (we stress that this condition is needed only for the proof provided below; the method remains applicable to more general cases).

Let us consider a solution of the form (6). Then, because of the orthonormality condition (13), we have $\langle \nabla f_\alpha, \nabla\tilde{\kappa}\rangle = c_\alpha = \langle \nabla f_\alpha, \tilde{\boldsymbol{u}}\rangle$, so that

$$\langle \nabla f_\alpha, \nabla\tilde{\kappa} - \tilde{\boldsymbol{u}}\rangle = \int_\Omega \nabla f_\alpha \cdot (\nabla\tilde{\kappa} - \tilde{\boldsymbol{u}})\, \mathrm{d}^2\theta = 0\,. \tag{A1}$$

Now we observe that the previous equation holds for any α: then it holds also for any linear combination $f = \sum_\alpha d_\alpha f_\alpha$ of $\{f_\alpha\}$. Thus

$$0 = \int_\Omega \nabla f \cdot (\nabla\tilde{\kappa} - \tilde{\boldsymbol{u}})\mathrm{d}^2\theta = -\int_\Omega f\nabla\cdot(\nabla\tilde{\kappa} - \tilde{\boldsymbol{u}})\, \mathrm{d}^2\theta + \int_{\partial\Omega} f(\nabla\tilde{\kappa} - \tilde{\boldsymbol{u}})\cdot\boldsymbol{n}\, \mathrm{d}\theta\,. \tag{A2}$$

In the last step we have integrated by parts ($\boldsymbol{n}$ is the unit vector orthogonal to the boundary $\partial\Omega$ of Ω).

We now use this equation to show that the chosen set of functions, described by Eq. (15), is *complete*, while, e.g., a similar set made of sine functions would not be complete. By the nature of the chosen set of functions we already know that we can properly represent any smooth function f. Using this property, we want to show that the two terms $\nabla\cdot(\nabla\tilde{\kappa} - \tilde{\boldsymbol{u}})$ and $(\nabla\tilde{\kappa} - \tilde{\boldsymbol{u}})\cdot\boldsymbol{n}$ entering in the r.h.s. of Eq. (A2) vanish on Ω and on $\partial\Omega$ respectively.

For the purpose, we observe that if cosines are used as set of functions, we can "build" any function f provided that the function has periodic derivatives on the boundary. In particular, if $A \subset \Omega$ is an arbitrary open subset of Ω, there is a function f that is positive on A and vanishes on $\Omega \setminus A$. Now suppose *per absurdum* that the solution obtained from the direct method does not satisfy Eq. (3), so that, e.g., $\nabla\cdot(\nabla\tilde{\kappa} - \tilde{\boldsymbol{u}}) > 0$ on a point $\boldsymbol{\theta}^* \in \Omega$. Then, for the sign persistence theorem, this quantity must be strictly positive in a neighborhood A of $\boldsymbol{\theta}^*$. However, if we take a function f which is positive on A and vanishes elsewhere, the rhs of Eq. (A2) will be positive, while the lhs vanishes, which is contradictory. This proves that Eq. (3) is verified by cosines.

In a similar manner, now that we have "disposed of" the first term, we observe that using cosines we can build a function f that vanishes everywhere on the boundary of $\partial\Omega$ except for a neighborhood. In other words, given an open subset $B \subset \partial\Omega$ of the boundary $\partial\Omega$, there is a function f that is positive on B and vanishes on $\partial\Omega \setminus B$. Note that this property would not be satisfied if the set of functions $\{f_\alpha\}$ were based on sines. Using a proof similar to the one given above, we obtain that $(\nabla\tilde{\kappa} - \tilde{\boldsymbol{u}})\cdot\boldsymbol{n}$ must vanish, thus leading to Eq. (4).

One might worry that, on the boundary, the chosen set of functions of Eq. (15) has zero derivative in the direction of $\boldsymbol{n}$, i.e. $\nabla f_{\alpha\beta}\cdot\boldsymbol{n} = 0$. This might suggest that the boundary condition (4) cannot be reproduced. In reality, although this is true *pointwise*, this does not affect

the convergence in $\mathcal{L}^2$ that is relevant for our problem (see Eqs. (7) and (12)). This important point is best illustrated by the following example.

Suppose that we measure a constant field $\tilde{\boldsymbol{u}}(\boldsymbol{\theta}) = (1, 0)$ on a square field Ω of side π. The obvious solution for $\tilde{\kappa}$ in this case is $\tilde{\kappa}(\theta) = \theta_1$. Suppose now that we try to use a set of functions made of sines. Then the corresponding coefficients $c_{\alpha\beta}$ would be proportional to the integrals $\alpha \int_0^\pi \theta_1 \cos\alpha\theta_1 \, \mathrm{d}\theta_1 \int_0^\pi \sin\beta\theta_2 \, \mathrm{d}\theta_2 = 0$. Hence, every coefficient $c_{\alpha\beta}$ vanishes. This proves that a set based on sines is not complete (condition (12) is not satisfied or, equivalently, a curl-free vector field $\tilde{\boldsymbol{u}}$ leads to a vanishing mass map). On the other hand, if we use the set (15), the coefficients $c_{\alpha\beta}$ do not vanish and the corresponding mass distribution is given by

$$\tilde{\kappa}(\boldsymbol{\theta}) = -\sum_{\alpha \text{ odd}} \frac{4}{\pi\alpha^2} \cos\alpha\theta_1 \ . \tag{A3}$$

This function can be shown to reduce to $\tilde{\kappa}(\boldsymbol{\theta}) = \theta_1 - \pi/2$.

References

Bartelmann M., 1995, A&A 303, 643
Brezis H., 1987, "Analyse fonctionnelle : théorie et application," Masson, Paris
Gelfand I.M., Fomin S.V., 1963, "Calculus of Variations," Prentice-Hall, Englewood Cliffs, NJ
Kaiser N., 1995, ApJ 493, L1
Lombardi M., Bertin G., 1998a, A&A 330, 791
Lombardi M., Bertin G., 1998b, A&A 335, 1
Lombardi M., Bertin G., 1999, A&A 342, 337
Schneider P., 1995, A&A 302, 639
Schneider P., Ehlers J., Falco E.E., 1992, "Gravitational Lenses," Springer-Verlag, Berlin
Seitz C., Schneider P., 1997, A&A 318, 687
Seitz S., Schneider P., 1996, A&A 305, 383
Seitz S., Schneider P., 1998, preprint astro-ph/9802051
Squires G., Kaiser N., 1996, ApJ 473, 65